

# **Quantitative CT analysis in ILD and use of artificial intelligence on imaging of ILD**

**Lucio Calandriello<sup>1</sup>, Tahreema Matin<sup>2</sup>, Helmut Prosch<sup>3</sup> and Joseph Jacob<sup>4,5</sup>**

<sup>1</sup>

Radiologia Diagnostica ed Interventistica Generale, Dipartimento di Diagnostica per Immagini,  
Radioterapia Oncologica ed Ematologia , Fondazione Policlinico Universitario A. Gemelli IRCCS,  
Roma, Italy

<sup>2</sup>Directorate of Education and Quality, Health Education England, London, UK.

<sup>3</sup>Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna,  
Vienna, Austria

<sup>4</sup>Department of Respiratory, University College London, UK

<sup>5</sup>Centre for Medical Image Computing, University College London, UK

Acknowledgements: Joseph Jacob was supported by a Wellcome Trust Clinical Research Career  
Development Fellowship 209553/Z/17/Z.

Corresponding author: Joseph Jacob; j.jacob@ucl.ac.uk

## **ABSTRACT**

Advances in computer technology over the past decade, particularly in the field of medical image analysis, have permitted the identification, characterization and quantitation of abnormalities that can be used to diagnose disease, or determine disease severity. On CT imaging performed in patients with interstitial lung disease (ILD), deep learning computer algorithms now demonstrate comparable performance with trained observers in the identification of a usual interstitial pneumonia pattern which is associated with a poor prognosis in several fibrosing ILDs.

Computer tools that quantify individual voxel-level CT features have also now come of age and can predict mortality with greater power than visual CT analysis scores. As these tools become more established, they have the potential to improve the sensitivity with which minor degrees of disease progression are identified. Currently pulmonary function tests are the gold standard measured used to assess clinical deterioration. However, the variation associated with pulmonary function measurements may mask the presence of small but genuine functional decline, which in the future could be confirmed by computer tools. The current review will describe the latest advances in quantitative CT analysis and deep learning as related to the ILDs and suggest potential future directions for this rapidly advancing field.

**KEY WORDS**

Interstitial lung disease

Quantitative computed tomography

Artificial Intelligence

Machine Learning

Lung fibrosis

Computed Tomography Imaging

Idiopathic pulmonary fibrosis

**DISCLOSURES**

Lucio Calandriello reports speaker fees from Roche and Boehringer Ingelheim unrelated to the current submission.

Joseph Jacob reports fees from Boehringer Ingelheim and Roche unrelated to the current submission.

Helmut Prosch reports research support from Boehringer Ingelheim and an institutional research cooperation with Siemens.

### 3.1 Introduction

High resolution computed tomography (CT) of the chest is central to the evaluation of patients with interstitial lung disease (ILD). One of the strengths of CT is its ability to identify signs of fibrosis. This allows the wide list of differential diagnoses of ILD, which include over a hundred differential entities, to five fibrotic ILDs (1). In the setting of fibrosing lung disease (FLD), a key initial assessment is the identification of a usual interstitial pneumonia (UIP) pattern. A UIP pattern classically presents with basal, subpleural and bilateral reticular opacities, bronchiectasis, and honeycombing (Figure 3.1). The identification of a UIP pattern in the right clinical context allows a diagnosis of idiopathic pulmonary fibrosis (IPF) to be made without the need for a lung biopsy (2). Thanks to its diagnostic ability, CT has been formally integrated into the multidisciplinary team discussion which has become the reference standard for the diagnosis of ILD (2-4).

CT also has a prognostic role in ILD evaluation. A typical UIP pattern on CT is associated with a worse prognosis in idiopathic disease, in chronic hypersensitivity pneumonitis and connective tissue disease-related ILD (5). In addition to the overall radiologic pattern, individual CT signs also demonstrate prognostic value. Extensive honeycombing, increased severity and extent of traction bronchiectasis and a high fibrosis score (the sum of honeycombing and reticulation extents) have been associated with an increased risk of death across several fibrotic lung diseases (6-10).

Despite its diagnostic and prognostic potential, CT interpretation has limitations. Firstly, CT evaluation is liable to significant interobserver variability. For example, the agreement between radiologists for the identification of honeycombing is surprisingly poor (11). Furthermore, the radiological classification of fibrotic interstitial lung disease according to patterns defined by ILD guidelines is subject to high interobserver variability (12). Importantly, subtle changes can be

challenging to capture in coarse visual scores (Figure 3.2). Accurate assessment of disease extent and severity on a patient's initial CT, would enable improved patient stratification and identification of individuals at higher risk of disease progression. This in turn could facilitate earlier intervention, for example selection for early lung transplant referral.

Whilst pulmonary function tests (PFTs) remain the conventional method used to monitor ILD progression, there are intrinsic disadvantages to assessing regional, heterogeneous lung changes using global measures of lung damage. Normal PFT ranges are wide and may mask small measurement deteriorations. A decline in patient health as evidenced by reduction of annual forced vital capacity (FVC) by 5-10% may not be accurately detected in the presence of variation due to patient effort/posture etc.

Some staging systems incorporating visual CT scores with physiologic parameters have been validated in ILD but due to limitations in observer agreement and low sensitivity are not routinely used in clinical practice or as drug trial end-points (13, 14). Moreover visual CT evaluation is not sensitive enough to capture short-term clinically useful changes (6, 11, 15, 16). As antifibrotic therapies improve, more sensitive biomarkers of disease progression will be required to capture their added therapeutic effect. There is also growing evidence that CTs may contain prognostic information that is not visually detectable but is amenable to computer-based quantification (17, 18). These limitations provide the rationale for computer-based CT analysis.

Our review will initially outline the CT acquisition techniques that can enable optimal performance of computer analysis tools. We then mimic a clinical workflow to our review by first describing computer analysis use in ILD diagnosis (where deep learning methods have been the mainstay)

before describing in turn the various quantitative tools that have been employed for measuring disease severity/extent in individual FLDs. We conclude by summarising the future of computer analysis in FLDs and outline the challenges that await on the journey towards regulatory agency approval.

### 3.2 CT Acquisition Technique

A standardized CT acquisition protocol is essential to improve the reproducibility of computer-based image analysis. Volumetric acquisitions with contiguous or overlapped thin-section reconstruction is mandatory, with a suggested slice thickness around 1 mm (19). The radiation dose ranges used to acquire ILD CTs varies widely and should be standardized (20). At present most CTs acquired clinically, and therefore analysed by computer tools, have utilized a dose of >1mSv as recommended by consensus guidelines. The study of the effect of Iterative reconstruction techniques on computer-based textural analysis is still in its infancy, and is increasing in importance as CTs are increasingly acquired with dose reduction techniques (19). Indeed studies comparing the effects of computer analysis on CTs acquired using normal, low and CXR equivalent radiation doses are keenly awaited to understand the dose range at which computer analysis will be constrained.

For image reconstruction, selection of a neutral kernel is optimal to avoid edge enhancement or image smoothing and most CT manufacturers provide bespoke neutral kernels (21). A sharper kernel can be analysed quantitatively if pre-processing image normalization methods are used (22). The depth of inspiration achieved during a scan is a major source of patient-related variation in lung attenuation measurements. Two possible approaches can be used to mitigate against inspiratory volume variability: respiratory gating to determine when the patient is at total lung capacity, or coaching of the patient to comply with standardized breathing instructions (21). In order to minimize differences between different scanners, as in case of trials where multiple centers using different scanners are involved, calibration of the scanner with test objects (phantoms) is suggested before and during data collection (20).

### **3.3 Artificial Intelligence**

Artificial intelligence (AI) is a general definition which encompasses several computer techniques meant to perform tasks usually performed by humans. The subfield of AI most commonly used in medicine is machine learning (ML) in which algorithms produce outputs based on patterns and features learned from input data (23, 24). Deep learning is a type of ML whose characteristic is its capability to automatically identify the most predictive features in a dataset and generate models for tasks from the raw data without the need for human programming (25). Deep learning can rely on several algorithms but those most commonly employed in medical imaging are Convolutional Neural Networks (CNN). The computer architecture at the base of a CNN is inspired by the human nervous system with clusters of artificial interconnected nodes replicating the interconnections of human brain neurons (23). The clusters of nodes in a CNN are usually organized in multiple layers: a layer that takes the input, numerous hidden layers that process the data and a layer which generates the output. Increasing the number of hidden layers allow the algorithm to deal with ever more complex tasks. When input data (i.e. images) are converted into digital data, a CNN can identify and extract imaging features that can be used to classify the data. A major advantage is the ability of a CNN to extract features that cannot be detected by human eyes (23).

In the setting of ILD, machine learning has already been used in several areas. These include: detecting the presence of ILD in patients with systemic sclerosis following analysis of PFTs(26), diagnosing IPF or a UIP pattern through genomic analysis (27, 28), quantifying lung fibrosis on CT images (29) and in providing automated classification of fibrotic ILD patterns on CT images (30).

#### **3.3.1 AI and ILD Diagnosis**



There are several reasons for implementing AI in the study of ILD. Visual radiological evaluation can be subject to low interobserver agreement when classifying HRCT signs of disease extent or severity. AI systems may also mitigate the potentially harmful consequences of human error in radiologic reporting (31).

One of the first attempts to use neural networks in assessing radiologic images of ILD dates back to 1990 when an artificial neural network was used to generate a differential diagnosis for ILD subtypes on chest radiographs (32). In that study the neural network was designed to distinguish between 9 ILDs on the basis of radiographic and clinical data. The results showed a diagnostic performance comparable to that of accredited radiologists and superior to that of radiology residents. In more recent years ML techniques have been applied to CTs of ILD patients to recognize and classify CT patterns, to classify fibrotic lung disease according to diagnostic guidelines and to quantify lung fibrosis.

The first paper where a CNN was used for the classification of ILD patterns on CT was published in 2016. In that study the CNN showed an accuracy of 85% for classifying 6 different pattern of disease (GGO, reticulation, consolidation, micronodules, honeycombing and a combination of GGO and reticulation) plus healthy tissue leading to 7 different imaging classes. These results were obtained using a dataset of 14696 image patches to train and test the software. The data was derived from 120 HRCTs where two radiologists had manually drawn regions of interest (ROIs) around every single pattern (33). The paper by Kim et al. (34) obtained an even higher accuracy for a CNN in classifying ROIs of HRCT images in ILD patients. In this paper an accuracy of up to 96% was achieved with a reduction in classification error between similar patterns (i.e. normal case vs emphysema and reticular opacity vs honeycombing) as the number of layers in the CNN increased.

Moreover, in this paper the CNN was compared to a more conventional classifier based on a support vector machine achieving an increase in accuracy of 6-9%. This is believed to result from the CNN being able to automatically extract features from data instead of using predefined engineered features. In both studies however, the need for manual identification of ROIs prior to classification limited the clinical applicability of the methods in real-world clinical practice.

To overcome this limitation and to obtain a more comprehensive evaluation of CT images of ILD patients, Gao et al. (35) proposed an “holistic” approach using a CNN for the classification of CT slices instead of ROIs. For comparison with previous studies the same CNN was also used for a ROI-based classification. The accuracy of the CNN for the identification of 6 tissue classes (normal, emphysema, ground glass opacity (GGO), fibrosis, micronodules and consolidation) was higher for the patch-based approach (87.9%) compared to holistic image classification (68.6%).

Only one paper has been published to date employing a deep learning algorithm to provide an automatic classification of fibrotic lung disease on CT (30) according to international guidelines (36). In the study by Walsh et al, the algorithm was trained, validated and tested using a database of 1157 CTs acquired in two institutions. The algorithm showed an accuracy of 76.4% for the classification of CTs into three categories provided by the 2011 consensus IPF diagnostic guidelines (UIP pattern, possible UIP pattern, inconsistent with UIP) (36). One of the main limitations of this approach was the requirement of labeling every CT used for algorithm training by a single radiologist whose interpretation bias would have potentially affected algorithm performance. To overcome this limitation, the software was further tested with a different population of 150 CTs where the reference standard was a consensus opinion of 91 expert thoracic radiologists. Here the software showed an accuracy of 73.3% outperforming 60 of the 91 thoracic radiologists.

Moreover, the interobserver agreement between the algorithm and the radiologists' majority opinion ( $\text{kw}=0.69$ ) was comparable to the median interobserver agreement between each thoracic radiologist and the radiologists' majority opinion ( $\text{kw}=0.67$ ).

### **3.4 Quantitative CT**

Quantitative CT (QCT) describes the numerous computer-based CT image analysis methods developed to measure structural lung damage in ILD. QCT methods offer improved objectivity, speed, reproducibility and the ability to scale up analysis to hundreds or thousands of CTs. QCT-derived metrics demonstrate potential as prognostic imaging biomarkers with reported utility in evaluating:

1. Disease severity at a single timepoint on a CT
2. Disease progression, quantifying longitudinal change on sequential CTs

Novel QCT imaging biomarkers are emerging through machine learning or deep learning techniques that are not appreciated by the human eye. New biomarkers represent features that have no morphological correlate or radiological descriptor. Quantification of overall lung histogram features, regional CT density changes, parenchymal textural features and other ancillary assessments by advanced algorithms have the potential to standardise and enhance the role of CT in ILD evaluation.

#### **3.4.1 Quantitative CT for assessing ILD severity**

Most QCT methods employ density and/or texture-based analysis of varying complexity. All methods require initial segmentation of the whole lungs from surrounding chest wall structures, which should ideally be an automated process requiring minimal manual correction.

Density mask techniques and evaluation of the lung-density histogram were amongst the first QCT methods developed to assess ILD severity (37). Quantification of emphysema using a low-density Hounsfield unit (HU) threshold is well-recognised. Similarly, high-density thresholds have been used to detect soft tissue density in the lung replacing air density as a result of pulmonary fibrosis(38). For normal lung, the CT density histogram peaks at approximately -800 HU and is left-skewed. The presence of fibrosis results in an increased mean lung density and decrease in the lung histogram kurtosis (peakedness) and skewness (asymmetry)(39). Yet correlations between such density histogram metrics and survival in patients with ILD have been generally poor(40-43). This reflects the challenge of capturing detailed regional information using a summary global density measure, often confounded by low density structures, for example traction bronchiectasis and honeycombing, as well as air trapping in hypersensitivity pneumonitis.

Newer QCT methods apply texture-based analysis to characterise, model and process imaging features at a voxel level (Table 3.1). These methods incorporate both morphological and density features. By simulating human visual perceptual and learning processes, texture-based algorithms attempt to determine the type of abnormality (for example emphysema vs honeycombing vs cysts), severity (fine vs course reticulation) and disease extent(29, 37, 40, 44-46). These complex QCT methods overlaying the QCT readout onto CT images, allowing regional ILD changes to be visualised by clinicians and patients(40).

### **3.4.2 Examples of QCT tools**

#### **3.4.2.1 Density histogram analysis**

Various CT density thresholds have been proposed for the assessment of ILD extent, including >-700 HU(47) and a range between -750 HU and -300 HU for the specific detection of GGO(48).

Inherent advantages of the density mask technique include universal applicability and convenience, since the method is based on easily appreciated CT Hounsfield unit values. CT density histogram measures have been found to be inferior to visual fibrosis scores and PFTs for survival prediction (43).

In a study comparing histogram features to textural analysis in a cohort of 95 patients with IPF, significant density histogram changes between baseline and 1-year follow-up CT including mean CT value of the whole lungs ( $P = 0.003$ ), skewness ( $P < 0.001$ ) and kurtosis ( $P < 0.001$ ) were consistent with disease progression(43). However, CT texture-derived total quantitative lung fibrosis (QLF) scores were reported to be better than CT density histogram measures for assessing longitudinal change in IPF. Whilst both QLF and kurtosis scores correlated with FVC at baseline, only QLF scores significantly correlated with longitudinal FVC change ( $p=0.57$ ;  $p<0.0001$ )(40).

#### **3.4.2.2 Computer Aided Lung Informatics for Pathology Evaluation and Rating (CALIPER)**

CALIPER was developed using thoracic radiologist consensus assessment of histopathologically confirmed training images(49). CALIPER characterises volumetric CT data using morphological and 3D histogram features within voxel volume units labelled according to the conventional radiological lexicon: normal lung, GGO, reticulation, subtypes of low-attenuation and honeycombing (Figure 3.3).

When CALIPER and visual CT scores were examined as predictors of survival in patients with IPF, CALIPER variables were more accurate in survival prediction than equivalent visual CT scores. CALIPER honeycombing extent independently predicted mortality (hazard ratio 1.18;  $p=0.002$ ) following adjustment for baseline disease severity using the composite physiologic index(17). Acknowledging that functional measures including DLco and CPI are stronger predictors of survival in patients with severe IPF, the greatest potential utility of QCT measures may be in patients with less extensive or early stage disease(50).

A unique attribute of CALIPER is the ability to quantify vessel-related structures (VRS). VRS, previously termed pulmonary vessel volume (PVV), corresponds to pulmonary vessels and associated structures, for example perivascular fibrosis which has no visually scored equivalent (Figure 3.4). CALIPER VRS was shown to be an independent predictor of mortality in IPF (hazard ratio 1.53;  $p<0.0001$ ) and an increase in VRS was found to be superior to various automated and visual CT scores in predicting mortality in IPF(17). Analysis of longitudinal changes in CALIPER CT parameters demonstrated that CALIPER VRS thresholds of change were the strongest CT predictor of mortality in independent discovery and validation cohorts (51). Thresholds of VRS change were also shown to be able to adjudicate marginal FVC declines of 5-9.9% in a composite staging system in patients with IPF(51).

Whilst the majority of research using QCT has been undertaken in IPF patients, several CALIPER studies have examined other fibrosing lung diseases. Studies in connective tissue disease-related ILD(52) and unclassifiable ILD(53) have emphasised that the prognostic signal from CALIPER variables, particularly VRS is not confined to patients with IPF. In hypersensitivity pneumonitis

patients(54), and patients with rheumatoid arthritis-related ILD(55), simple binary VRS thresholds have shown the ability to identify patients with an IPF-like outcome, regardless of a patients starting baseline disease severity (Figure 3.4).

#### **3.4.2.3 Automated Quantification System (AQS)**

AQS, developed in Seoul, Korea quantifies six regional patterns on CT: normal, GGO, reticular opacity (RO), honeycombing, emphysema and consolidation. An AQS fibrotic score is calculated from the sum of RO and honeycombing extent. The texture-based AQS model was trained using multiple sampled regions from CT images acquired using different CT scanners by a thoracic radiologist with 10 years of thoracic CT experience(45).

AQS “RO” showed the strongest correlation with FVC ( $r=-0.435$ ;  $p<0.001$ ) and AQS fibrosis score was most strongly correlated with DLco ( $r=-0.506$ ;  $p<0.001$ ). Multivariate analysis identified RO as the only independent predictor for FVC decline ( $p=0.021$ ; adjusted odds ratio 1.047). Receiver operating characteristic curve analysis found that an RO cut-off value of 22.05% (sensitivity 50.0%; specificity 81.4%; negative predictive value 89.1%) represents a level below which FVC is likely to be stable at 1-year follow up in IPF patients (45).

#### **3.4.2.4 Adaptive multiple feature method (AMFM)**

AMFM is a texture-based analysis tool designed to recognise the following CT patterns: normal lung, GGO, ground-glass reticulation (GGR), emphysema and honeycombing. The AMFM tool employs 26 mathematical features to describe regional density patterns combined with a Bayesian classifier to quantify volumes of the various radiological patterns. The method was trained on annotated volumes of interest from CTs by three expert radiologists(46). Sensitivity and specificity

for automatic identification by three-dimensional AMFM was reported as 100% for emphysema, 95% for consolidation and 97% for honeycombing(56).

Ground-glass reticulation (GGR) was the AMFM feature that showed the strongest correlation with equivalent visually-scored CT features ( $r=-0.60$ ;  $p<0.0001$ )(46). Retrospective application of AMFM to a large cohort of patients with IPF demonstrated that GGR was independently associated with disease progression(46). In 105 patients with IPF with multiple follow-up CTs (median interval between baseline and final CT was 63.9 weeks), increasing GGR measured by AMFM software moderately correlated with decreasing FVC ( $r=-0.25$ ;  $p=0.01$ )(46). Similarly, a three-expert average of visually-scored GGR also moderately correlated with FVC change ( $r=-0.30$ ;  $p=0.002$ ). Visual and AMFM change assessment for GGR was moderately correlated ( $r = 0.47$ ,  $P < 0.0001$ ).

#### **3.4.2.5 Quantitative Lung Fibrosis**

Quantitative Lung Fibrosis (QLF) and Quantitative Interstitial Lung Disease (QILD) are two scores provided by an automated CAD-based system developed for assessing ILD using texture features (Figure 3.5 and 3.6)(57). A Support Vector Machine (SVM) was used to select features to classify fibrotic reticulations, ground glass opacities, honeycombing and normal lung parenchyma. The texture-based scores were obtained using a 5 step process: 1) denoising CT images, 2) sampling each pixel from a 4x4 grid within the segmented lung, 3) extracting texture features using characteristics of grid intensities, 4) using a SVM to classify pixels, 5) expressing the scores as a percentage of lung involvement or volume of affected lung tissue (40). Denoising the CT images allowed normalization of differences in acquisition parameters across longitudinal CTs (40).



The automated system was initially used to classify disease patterns in selected ROIs (58) but was extended to quantify lung fibrosis across the entire lung (57). The system produced objective and reproducible quantification of lung fibrosis in patients with IPF and assessed changes in fibrosis extent in scleroderma patients undergoing treatment. In IPF patients the QLF score demonstrated significant correlations between change in QLF values across longitudinal CTs and change in pulmonary function tests (40). Notably the improved detection of treatment effects within a given sample size was shown to allow an approximately 50% reduction in the study sample size whilst maintaining the same study power (59).

#### **3.4.2.6 Data-Driven Textural Analysis**

Data-Driven Textural Analysis is a computer algorithm which uses an unsupervised feature-learning process, based on a CNN, to select features to train a SVM classifier (29). This unsupervised process uses a huge number of image patches to produce a set of basic elements (dictionary) which can be used to classify other image regions. The system was then trained as a classifier using ROIs labeled by a radiologist as normal or fibrotic lung. In this way the SVM is trained to classify new image regions based on the established dictionary and categorize them as either normal lung or fibrotic lung. This approach is different from other more traditional methods where features are manually defined. The main advantage of the system is its use of more discriminative features for classifier training which makes automatic classification of lung patterns on CT more effective (Figure 3.7).

The DTA approach has been used on a population of 280 IPF patients, with 72 of these undergoing follow-up CT. DTA fibrosis score, calculated as the number of fibrotic ROIs divided by the total number of ROIs sampled, showed significant correlations with semiquantitative visual scores ( $r=-$

0.50;  $p < 0.001$ ) and PFTs (FVC ( $r = -0.60$ ,  $p < 0.001$ ), DLco ( $r = -0.68$ ,  $p < 0.001$ )) at baseline. In patients with follow-up CT imaging (Figure 3.7), change in DTA fibrosis score showed moderate significant correlations with change in FVC ( $\rho = -0.41$ ;  $p < 0.001$ ) and DLco ( $\rho = -0.40$ ;  $p < 0.001$ ). These findings suggest that DTA can provide an objective measurement of baseline disease severity and disease progression (29). Further work comparing longitudinal DTA change to physiological and patient-reported outcome measures have shown that the minimal clinically important difference in DTA fibrosis score is 3.4% of the lung (60).

#### **3.4.2.7 Functional Respiratory Imaging**

Functional respiratory imaging (FRI) is a post-processing application that uses volumetric CT images acquired at functional respiratory capacity (FRC) and total lung capacity (TLC).

Computational fluid dynamics (CFD) are used to obtain structural and functional parameters from CT imaging. CTs are acquired using respiratory gating to guarantee correct and repeatable lung volumes. FRI calculates lobar volumes, airway and vessel volumes, airway resistance and airway radius (61, 62). FRI is the only QCT tool to consider airway volumes as potential prognostic variables. FRI has been validated in obstructive pulmonary disease (19) with low variability in airway and vessel volume measurements and measures of airways resistance (63, 64).

In IPF patients, significant correlations were found between declining FVC and declining lung volumes, increasing fibrotic tissue (calculated as areas with HU between -600 and 600) and increases in specific airway radius as measured by FRI. The correlation between FVC decline and an increase in airway volumes is likely to relate to the development of traction bronchiectasis, with the correlation more pronounced in the lower lobes. The utility of measuring changes in airway radius to predict disease severity and progression may be accentuated in patients with

early disease where FVC measurements can still be normal. The ability to detect regional changes in airway or lobar volumes might allow a reduction in sample sizes needed for IPF drug trials to detect disease progression or treatment effects (61).

### **3.5 Future Avenues and Challenges**

As our understanding of FLD and computer-based image analysis evolves together, new areas of research targeting unmet clinical needs are emerging. It is well established that in cases where a diagnosis may not be clear, observing the behavior of a disease can provide a provisional working diagnosis for a patient (65). More recently the concept of the progressive fibrotic phenotype (66) has emerged to define a patient cohort that cuts across diagnoses, where administration of anti-fibrotic medication may be beneficial. Identifying this patient group using automated means could be achieved by analysing time-series data, modelled on disease progression/survival independent of their actual diagnosis. Algorithms could then identify CT features in patients with a non-idiopathic fibrosing lung disease that suggest an IPF-like outcome. Using the same diagnosis-agnostic approach, disease classification may be based on likely prognosis, which might better inform management strategies. Accurate determination of disease trajectory is undoubtedly aided by identifying sensitive, reproducible biomarkers of disease progression. In this regard, improvements in the measurements of airway metrics such as airway volume and tortuosity across the fibrosing lung diseases hold great promise.

The characterization of interstitial lung abnormalities which may represent an early stage of fibrosing lung disease is another area where large volumes of data should provide opportunities for deep learning algorithms. The advent of lung cancer screening programs in several countries is generating imaging and outcome data in a population enriched in their propensity to develop

fibrosing lung disease. It is here that patterns identified by a computer, but which have no visual equivalent might have the most promise. Given the huge numbers of patients that are imaged in lung cancer screening programs, some form of automated triage of chest CT imaging could also rationalize referrals to interstitial lung disease units. Yet the sheer heterogeneity of the CT imaging in screening programs will need to be addressed to avoid false positive and false negative identification of disease. Heterogeneity may result from technical parameters such as differences in scanner manufacturer, model, reconstruction algorithm and radiation dose between studies. Additionally, patient factors such as variation in inspiratory effort, concurrent illness, patient weight and position (prone versus supine) differences can affect automated CT interpretation. Deep learning algorithms are now being employed to harmonize CTs acquired using different techniques in an attempt to limit interscan measurement variability.

Identifying and quantifying the complex imaging features seen on FLD CTs could potentially be improved by capturing volumetric information with three-dimensional deep neural networks. Though image interpretation could be improved when compared to two-dimensional neural networks, three-dimensional deep networks typically require far more processing power and much larger datasets for model training. ILD databases today generally number hundreds of patients which might be too small for robust image classification by three-dimensional networks. Though existing limitations in computer hardware are likely to be overcome in the near future, the constraint of limited imaging data in the field of FLD will be a harder challenge to overcome. It is also becoming apparent that evaluating imaging alone will not fully harness the power of computational analysis. FLD datasets of the future are likely to be multimodal comprising electronic health records, imaging, genetic and biomarker information. Interpreting this information will improve our understanding of disease pathogenesis but will require a new kind of

clinical team where bioinformaticians, geneticists, statisticians and computer scientists might work together in truly multidisciplinary collaborations that cut across scientific fields.

Finally, the question of trust is becoming increasingly pertinent to the study of computer analysis. As with self-driving cars, in the sphere of medical image analysis, we hold computers to a higher standard than we might expect of humans. And this is because the potential implications of an algorithm being blind to an imaging abnormality due to deficient training data is real and concerning. Standards are now being proposed by which we should evaluate computer analysis of medical imaging and these will be our guide in the years to come (67). A clear and fundamental imperative is that algorithms are rigorously validated across multiple diverse datasets prior to widespread use, to rule out overfitting of an algorithm to its training datasets. Furthermore, as algorithms are heavily influenced by the data used for training, it is important to ensure that patients subgroups are adequately represented e.g. by actively considering the distribution of demographics such as gender, ethnicity and socioeconomic status in training datasets. Similarly, in relation to FLD, algorithms trained on western European or North American populations may not transpose easily to developing world settings where biomass-fuel exposure and tuberculosis-related lung damage may affect image interpretation. This is before considering the well-documented regional variations in prevalence of various FLDs around the world(68).

### **3.6 Conclusions**

Over a relatively short space of time, computer analysis of ILD imaging has found an important niche in the evaluation of patients with fibrosing lung disease. Though primarily research tools at present, as the field evolves and matures and acquires approval by regulatory agencies, their incorporation into clinical radiology workflows can be envisioned. The main opportunities lie in

quantifying disease severity at baseline, disease progression over time and characterizing interstitial lung abnormalities, particularly in lung cancer screening populations. This may herald a change in the vocabulary used to describe lung fibrosis, which has remained quite consistent over 20 years. Challenges remain, particularly in relation to the availability of comprehensive datasets for analysis. Yet given the complexities and challenges inherent to visual ILD evaluation, the integration of objective measures of disease subtype, severity and progression should be embraced.

## REFERENCES

1. Jacob J, Hansell DM. HRCT of fibrosing lung disease. *Respirology*. 2015;20(6):859-72.
2. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med*. 2018;198(5):e44-e68.
3. American Thoracic S, European Respiratory S. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. *Am J Respir Crit Care Med*. 2002;165(2):277-304.
4. Travis WD, Costabel U, Hansell DM, King TE, Jr., Lynch DA, Nicholson AG, et al. An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med*. 2013;188(6):733-48.
5. Flaherty KR, Thwaite EL, Kazerooni EA, Gross BH, Toews GB, Colby TV, et al. Radiological versus histological diagnosis in UIP and NSIP: survival implications. *Thorax*. 2003;58(2):143-8.
6. Lynch DA, Godwin JD, Safrin S, Starko KM, Hormel P, Brown KK, et al. High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *Am J Respir Crit Care Med*. 2005;172(4):488-93.
7. Walsh SL, Sverzellati N, Devaraj A, Keir GJ, Wells AU, Hansell DM. Connective tissue disease related fibrotic lung disease: high resolution computed tomographic and pulmonary function indices as prognostic determinants. *Thorax*. 2014;69(3):216-22.
8. Walsh SL, Sverzellati N, Devaraj A, Wells AU, Hansell DM. Chronic hypersensitivity pneumonitis: high resolution computed tomography patterns and pulmonary function indices as prognostic determinants. *European radiology*. 2012;22(8):1672-9.
9. Edey AJ, Devaraj AA, Barker RP, Nicholson AG, Wells AU, Hansell DM. Fibrotic idiopathic interstitial pneumonias: HRCT findings that predict mortality. *European radiology*. 2011;21(8):1586-93.
10. Lee HY, Lee KS, Jeong YJ, Hwang JH, Kim HJ, Chung MP, et al. High-resolution CT findings in fibrotic idiopathic interstitial pneumonias with little honeycombing: serial changes and prognostic implications. *AJR American journal of roentgenology*. 2012;199(5):982-9.
11. Watadani T, Sakai F, Johkoh T, Noma S, Akira M, Fujimoto K, et al. Interobserver variability in the CT assessment of honeycombing in the lungs. *Radiology*. 2013;266(3):936-44.
12. Walsh SL, Calandriello L, Sverzellati N, Wells AU, Hansell DM, Consort UIPO. Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax*. 2016;71(1):45-51.
13. Walsh SL, Wells AU, Sverzellati N, Keir GJ, Calandriello L, Antoniou KM, et al. An integrated clinicoradiological staging system for pulmonary sarcoidosis: a case-cohort study. *Lancet Respir Med*. 2014;2(2):123-30.
14. Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ, Maher TM, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med*. 2008;177(11):1248-54.
15. Walsh SL, Hansell DM. High-resolution CT of interstitial lung disease: a continuous evolution. *Semin Respir Crit Care Med*. 2014;35(1):129-44.
16. Elicker BM, Kallianos KG, Henry TS. The role of high-resolution computed tomography in the follow-up of diffuse lung disease: Number 2 in the Series "Radiology" Edited by Nicola Sverzellati and Sujal Desai. *European respiratory review : an official journal of the European Respiratory Society*. 2017;26(144).
17. Jacob J, Bartholmai BJ, Rajagopalan S, Kokosi M, Nair A, Karwoski R, et al. Mortality prediction in idiopathic pulmonary fibrosis: evaluation of computer-based CT analysis with conventional severity measures. *Eur Respir J*. 2017;49(1).
18. Jacob J, Bartholmai BJ, Rajagopalan S, Brun AL, Egashira R, Karwoski R, et al. Evaluation of computer-based computer tomography stratification against outcome models in connective tissue disease-related interstitial lung disease: a patient outcome study. *BMC medicine*. 2016;14(1):190.
19. Wu X, Kim GH, Salisbury ML, Barber D, Bartholmai BJ, Brown KK, et al. Computed Tomographic Biomarkers in Idiopathic Pulmonary Fibrosis. The Future of Quantitative Analysis. *Am J Respir Crit Care Med*. 2019;199(1):12-21.

20. Sieren JP, Newell JD, Jr., Barr RG, Bleecker ER, Burnette N, Carretta EE, et al. SPIROMICS Protocol for Multicenter Quantitative Computed Tomography to Phenotype the Lungs. *Am J Respir Crit Care Med*. 2016;194(7):794-806.
21. Newell JD, Jr., Sieren J, Hoffman EA. Development of quantitative computed tomography lung protocols. *J Thorac Imaging*. 2013;28(5):266-71.
22. Gallardo-Estrella L, Lynch DA, Prokop M, Stinson D, Zach J, Judy PF, et al. Normalizing computed tomography data reconstructed with different filter kernels: effect on emphysema quantification. *European radiology*. 2016;26(2):478-86.
23. Dreyer KJ, Geis JR. When Machines Think: Radiology's Next Frontier. *Radiology*. 2017;285(3):713-8.
24. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, et al. Deep Learning: A Primer for Radiologists. *Radiographics*. 2017;37(7):2113-31.
25. Lee SM, Seo JB, Yun J, Cho YH, Vogel-Claussen J, Schiebler ML, et al. Deep Learning Applications in Chest Radiography and Computed Tomography: Current State of the Art. *J Thorac Imaging*. 2019;34(2):75-85.
26. Le-Dong N-N, Hua-Huy T, Nguyen-Ngoc H-M, Dinh-Xuan A-T. Applying machine learning and pulmonary function data to detect interstitial lung disease in systemic sclerosis. *Eur Respir J*. 2017;50(suppl 61):OA3438.
27. Pankratz DG, Choi Y, Imtiaz U, Fedorowicz GM, Anderson JD, Colby TV, et al. Usual Interstitial Pneumonia Can Be Detected in Transbronchial Biopsies Using Machine Learning. *Annals of the American Thoracic Society*. 2017;14(11):1646-54.
28. Kim SY, Diggans J, Pankratz D, Huang J, Pagan M, Sindy N, et al. Classification of usual interstitial pneumonia in patients with interstitial lung disease: assessment of a machine learning approach using high-dimensional transcriptional data. *The Lancet Respiratory medicine*. 2015;3(6):473-82.
29. Humphries SM, Yagihashi K, Huckleberry J, Rho BH, Schroeder JD, Strand M, et al. Idiopathic Pulmonary Fibrosis: Data-driven Textural Analysis of Extent of Fibrosis at Baseline and 15-Month Follow-up. *Radiology*. 2017;285(1):270-8.
30. Walsh SLF, Calandriello L, Silva M, Sverzellati N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study. *The Lancet Respiratory medicine*. 2018;6(11):837-45.
31. Fazal MI, Patel ME, Tye J, Gupta Y. The past, present and future role of artificial intelligence in imaging. *European journal of radiology*. 2018;105:246-50.
32. Asada N, Doi K, MacMahon H, Montner SM, Giger ML, Abe C, et al. Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study. *Radiology*. 1990;177(3):857-60.
33. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. *IEEE transactions on medical imaging*. 2016;35(5):1207-16.
34. Kim GB, Jung KH, Lee Y, Kim HJ, Kim N, Jun S, et al. Comparison of Shallow and Deep Learning Methods on Classifying the Regional Pattern of Diffuse Lung Disease. *Journal of digital imaging*. 2018;31(4):415-24.
35. Gao M, Bagci U, Lu L, Wu A, Buty M, Shin HC, et al. Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Computer methods in biomechanics and biomedical engineering Imaging & visualization*. 2018;6(1):1-6.
36. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med*. 2011;183(6):788-824.
37. Ohkubo H, Nakagawa H, Niimi A. Computer-based quantitative computed tomography image analysis in idiopathic pulmonary fibrosis: A mini review. *Respir Investig*. 2018;56(1):5-13.
38. Ohkubo H, Kanemitsu Y, Uemura T, Takakuwa O, Takemura M, Maeno K, et al. Normal Lung Quantification in Usual Interstitial Pneumonia Pattern: The Impact of Threshold-based Volumetric CT Analysis for the Staging of Idiopathic Pulmonary Fibrosis. *PLoS One*. 2016;11(3):e0152505.



39. Hartley PG, Galvin JR, Hunninghake GW, Merchant JA, Yagla SJ, Speakman SB, et al. High-resolution CT-derived measures of lung density are valid indexes of interstitial lung disease. *J Appl Physiol* (1985). 1994;76(1):271-7.
40. Kim HJ, Brown MS, Chong D, Gjertson DW, Lu P, Kim HJ, et al. Comparison of the quantitative CT imaging biomarkers of idiopathic pulmonary fibrosis at baseline and early change with an interval of 7 months. *Academic radiology*. 2015;22(1):70-80.
41. Ash SY, Harmouche R, Vallejo DL, Villalba JA, Ostridge K, Gunville R, et al. Densitometric and local histogram based analysis of computed tomography images in patients with idiopathic pulmonary fibrosis. *Respir Res*. 2017;18(1):45.
42. Best AC, Lynch AM, Bozic CM, Miller D, Grunwald GK, Lynch DA. Quantitative CT indexes in idiopathic pulmonary fibrosis: relationship with physiologic impairment. *Radiology*. 2003;228(2):407-14.
43. Best AC, Meng J, Lynch AM, Bozic CM, Miller D, Grunwald GK, et al. Idiopathic Pulmonary Fibrosis: Physiologic Tests, Quantitative CT Indexes, and CT Visual Scores as Predictors of Mortality. *Radiology*. 2008;246(3):935-40.
44. Maldonado F, Moua T, Rajagopalan S, Karwoski RA, Raghunath S, Decker PA, et al. Automated quantification of radiological patterns predicts survival in idiopathic pulmonary fibrosis. *Eur Respir J*. 2014;43(1):204-12.
45. Park HJ, Lee SM, Song JW, Lee SM, Oh SY, Kim N, et al. Texture-Based Automated Quantitative Assessment of Regional Patterns on Initial CT in Patients With Idiopathic Pulmonary Fibrosis: Relationship to Decline in Forced Vital Capacity. *AJR Am J Roentgenol*. 2016;207(5):976-83.
46. Salisbury ML, Lynch DA, van Beek EJ, Kazerooni EA, Guo J, Xia M, et al. Idiopathic Pulmonary Fibrosis: The Association between the Adaptive Multiple Features Method and Fibrosis Outcomes. *Am J Respir Crit Care Med*. 2017;195(7):921-9.
47. Shin KE, Chung MJ, Jung MP, Choe BK, Lee KS. Quantitative computed tomographic indexes in diffuse interstitial lung disease: correlation with physiologic tests and computed tomography visual scores. *J Comput Assist Tomogr*. 2011;35(2):266-71.
48. Kauczor HU, Heitmann K, Heussel CP, Marwede D, Uthmann T, Thelen M. Automatic detection and quantification of ground-glass opacities on high-resolution CT using multiple neural networks: comparison with a density mask. *AJR Am J Roentgenol*. 2000;175(5):1329-34.
49. Jacob J, Bartholmai BJ, Rajagopalan S, Kokosi M, Nair A, Karwoski R, et al. Automated Quantitative Computed Tomography Versus Visual Computed Tomography Scoring in Idiopathic Pulmonary Fibrosis: Validation Against Pulmonary Function. *J Thorac Imaging*. 2016;31(5):304-11.
50. Jacob J, Bartholmai BJ, Rajagopalan S, van Moorsel CHM, van Es HW, van Beek FT, et al. Predicting Outcomes in Idiopathic Pulmonary Fibrosis Using Automated Computed Tomographic Analysis. *Am J Respir Crit Care Med*. 2018;198(6):767-76.
51. Jacob J, Bartholmai BJ, van Moorsel CHM, Rajagopalan S, Devaraj A, van Es HW, et al. Longitudinal prediction of outcome in idiopathic pulmonary fibrosis using automated CT analysis. *bioRxiv*. 2018:493544.
52. Jacob J, Bartholmai BJ, Rajagopalan S, Brun AL, Egashira R, Karwoski R, et al. Evaluation of computer-based computer tomography stratification against outcome models in connective tissue disease-related interstitial lung disease: a patient outcome study. *BMC Med*. 2016;14:190.
53. Jacob J, Bartholmai BJ, Rajagopalan S, Egashira R, Brun AL, Kokosi M, et al. Unclassifiable-interstitial lung disease: Outcome prediction using CT and functional indices. *Respir Med*. 2017;130:43-51.
54. Jacob J, Bartholmai BJ, Egashira R, Brun AL, Rajagopalan S, Karwoski R, et al. Chronic hypersensitivity pneumonitis: identification of key prognostic determinants using automated CT analysis. *BMC Pulm Med*. 2017;17(1):81.
55. Jacob J, Hirani N, van Moorsel CHM, Rajagopalan S, Murchison JT, van Es HW, et al. Predicting outcomes in rheumatoid arthritis related interstitial lung disease. *Eur Respir J*. 2018:1800869.
56. Xu Y, van Beek EJ, Hwanjo Y, Guo J, McLennan G, Hoffman EA. Computer-aided classification of interstitial lung diseases via MDCT: 3D adaptive multiple feature method (3D AMFM). *Acad Radiol*. 2006;13(8):969-78.
57. Kim HG, Tashkin DP, Clements PJ, Li G, Brown MS, Elashoff R, et al. A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. *Clinical and experimental rheumatology*. 2010;28(5 Suppl 62):S26-35.

58. Kim HJ, Li G, Gjertson D, Elashoff R, Shah SK, Ochs R, et al. Classification of parenchymal abnormality in scleroderma lung using a novel approach to denoise images collected via a multicenter study. *Academic radiology*. 2008;15(8):1004-16.
59. Kim HJ, Brown MS, Elashoff R, Li G, Gjertson DW, Lynch DA, et al. Quantitative texture-based assessment of one-year changes in fibrotic reticular patterns on HRCT in scleroderma lung disease treated with oral cyclophosphamide. *European radiology*. 2011;21(12):2455-65.
60. Humphries SM, Swigris JJ, Brown KK, Strand M, Gong Q, Sundry JS, et al. Quantitative high-resolution computed tomography fibrosis score: performance characteristics in idiopathic pulmonary fibrosis. *Eur Respir J*. 2018;52(3).
61. Clukers J, Lanclus M, Mignot B, Van Holsbeke C, Roseman J, Porter S, et al. Quantitative CT analysis using functional imaging is superior in describing disease progression in idiopathic pulmonary fibrosis compared to forced vital capacity. *Respiratory research*. 2018;19(1):213.
62. Hajian B, De Backer J, Vos W, Van Holsbeke C, Clukers J, De Backer W. Functional respiratory imaging (FRI) for optimizing therapy development and patient care. *Expert review of respiratory medicine*. 2016;10(2):193-206.
63. Hajian B, De Backer J, Vos W, Van Holsbeke C, Ferreira F, Quinn DA, et al. Pulmonary vascular effects of pulsed inhaled nitric oxide in COPD patients with pulmonary hypertension. *International journal of chronic obstructive pulmonary disease*. 2016;11:1533-41.
64. De Backer J, Van Holsbeke C, Vos W, Vinchurkar S, Dorinsky P, Rebello J, et al. Assessment of lung deposition and analysis of the effect of fluticasone/salmeterol hydrofluoroalkane (HFA) pressurized metered dose inhaler (pMDI) in stable persistent asthma patients using functional respiratory imaging. *Expert review of respiratory medicine*. 2016;10(8):927-33.
65. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med*. 2018;198(5):e44-e68.
66. Wells AU, Brown KK, Flaherty KR, Kolb M, Thannickal Victor J. What's in a name? That which we call IPF, by any other name would act the same. *Eur Respir J*. 2018;51(5):1800692.
67. Allen B, Jr., Seltzer SE, Langlotz CP, Dreyer KP, Summers RM, Petrick N, et al. A Road Map for Translational Research on Artificial Intelligence in Medical Imaging: From the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop. *Journal of the American College of Radiology*. 2019;16(9):1179-89.
68. Singh S, Collins BF, Sharma BB, Joshi JM, Talwar D, Katiyar S, et al. Interstitial Lung Disease in India. Results of a Prospective Registry. *Am J Respir Crit Care Med*. 2017;195(6):801-13.

Quantitative computer tool	CT pattern identification
CALIPER (Computer-Aided Lung Informatics for Pathology Evaluation and Rating)	Measures low attenuation areas, ground glass opacities, reticulation, honeycombing, normal lung and vessel-related structures. Features identified in 15x15x15 voxel volume units on volumetric non-contrast enhanced CT images.
AQS (Automated Quantification System)	Measures low attenuation areas, ground glass opacities, reticular opacity, honeycombing, normal lung, emphysema and consolidation on non-contiguous, non-contrast enhanced CT images.
AMFM (Adaptive Multiple Features Method)	Measures ground glass opacities, ground-glass reticulation, honeycombing, normal lung and emphysema on volumetric non-contrast enhanced CT images.
QLF (Quantitative Lung Fibrosis)	Measures lung fibrosis (reticulation) or interstitial lung disease (sum of ground glass opacities, reticulation and honeycombing). Evaluates non-contiguous, non-contrast enhanced CT images.
DTA (Data-driven Textural Analysis)	Discriminates tissue as either fibrotic or normal. Fibrotic tissue includes reticulation, traction bronchiectasis and honeycombing. Evaluates non-contiguous, non-contrast enhanced CT images.
FRI (Functional Residual Imaging)	Measures lung, lobar, vessel and airway volumes and fibrotic and emphysematous volumes. Evaluates volumetric, non-contrast and contrast enhanced CT images.

Table 3.1 Quantitative computer tools that evaluate lung damage on computed tomography imaging

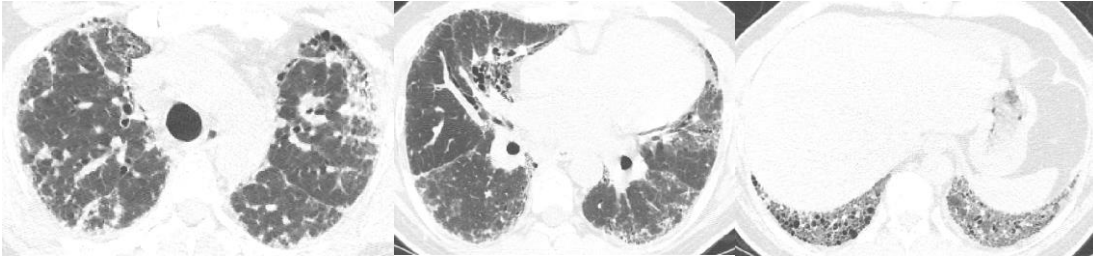


Figure 3.1 Axial CT images taken at the upper, middle and lower zones of the lungs demonstrating a usual interstitial pneumonia pattern on CT in a 49-year-old male 3-pack-year ex-smoker. The usual interstitial pneumonia pattern is characterized by honeycomb cysts and traction bronchiectasis in a predominantly basal and peripheral distribution within the lungs. In the presence of idiopathic disease, the patient was diagnosed with idiopathic pulmonary fibrosis following multi-disciplinary team discussion.

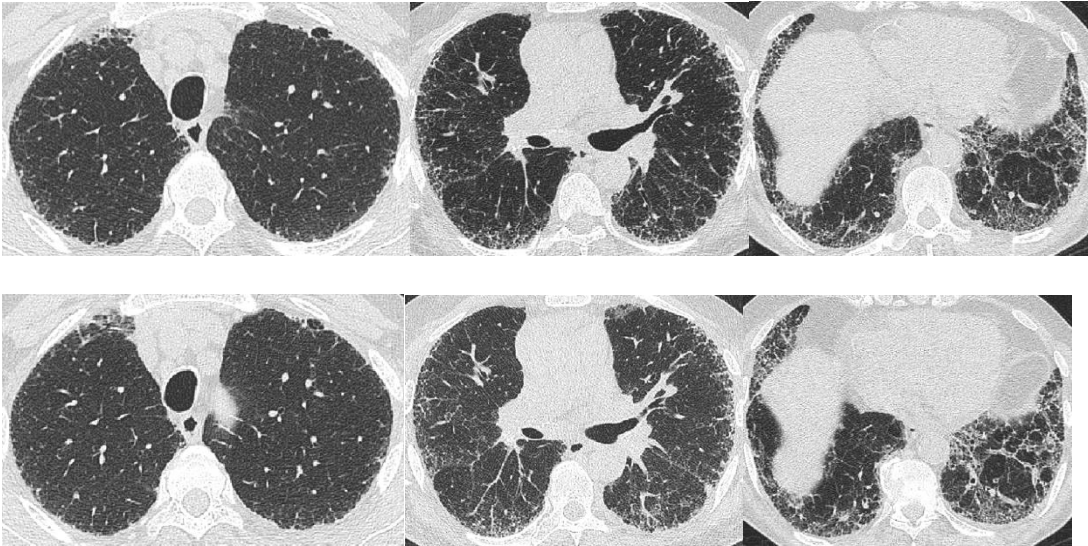


Figure 3.2 Changes on longitudinal CT images performed 15 months apart in a 77-year-old male, never-smoker diagnosed with idiopathic pulmonary fibrosis. Images have been taken at corresponding points in the upper, middle and lower zones of the lungs. Using traditional visual CT scoring methods quantifying total interstitial disease (ILD) extent to the nearest 5% on each CT timepoint, an increase in total ILD extent of more than 5% was not demonstrated. Yet the patient demonstrated clinical deterioration. Whilst there is subtle alteration in the character of the fibrotic regions, it did not equate to a measurable difference in total ILD extent for the whole lung.

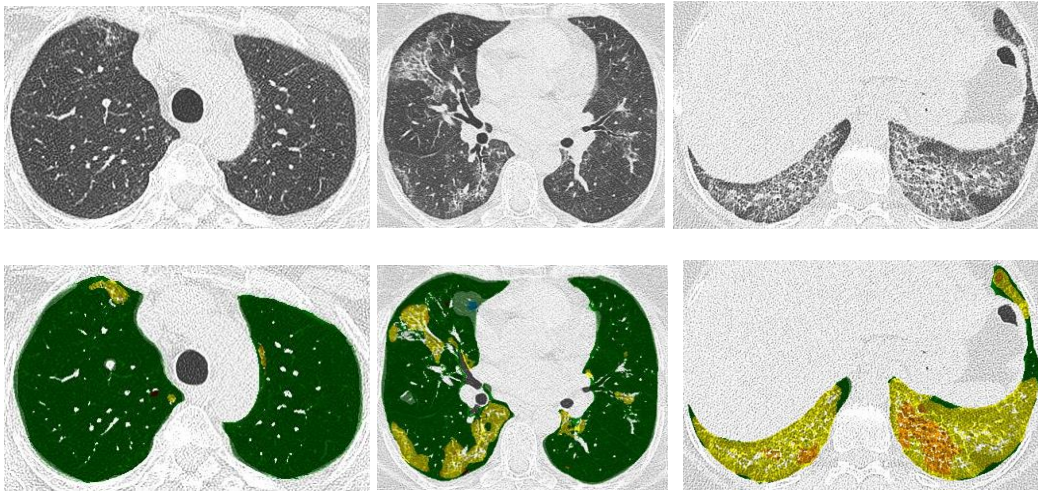


Figure 3.3 Axial CT images in the upper, middle and lower zones of the lungs (top row) in a 54-year-old female never-smoker diagnosed with rheumatoid arthritis-related interstitial lung disease. Corresponding images on the bottom row demonstrate parenchymal quantitation by CALIPER. The images demonstrate fibrosis characterized by reticulation, traction bronchiectasis and architectural distortion with a lower zone predominance. (Dark green=normal lung, light green and blue=reduced attenuation lung, yellow=ground glass opacity, orange=reticulation).



Figure 3.4 Axial CT image of a 77-year-old male ex-smoker with rheumatoid arthritis-related interstitial lung disease. Fibrosis with peripheral layered honeycombing cysts are visible in the midzones of the lung. Vessel-related structures (pulmonary arteries and veins and related fibrosis) quantified by CALIPER, a quantitative CT tool, are highlighted in red. In patients with rheumatoid arthritis-related interstitial lung disease, vessel-related structures comprising more than 4.4% of the lung have been shown to predict an idiopathic pulmonary fibrosis-like outcome, regardless of a patients' baseline disease extent or severity.



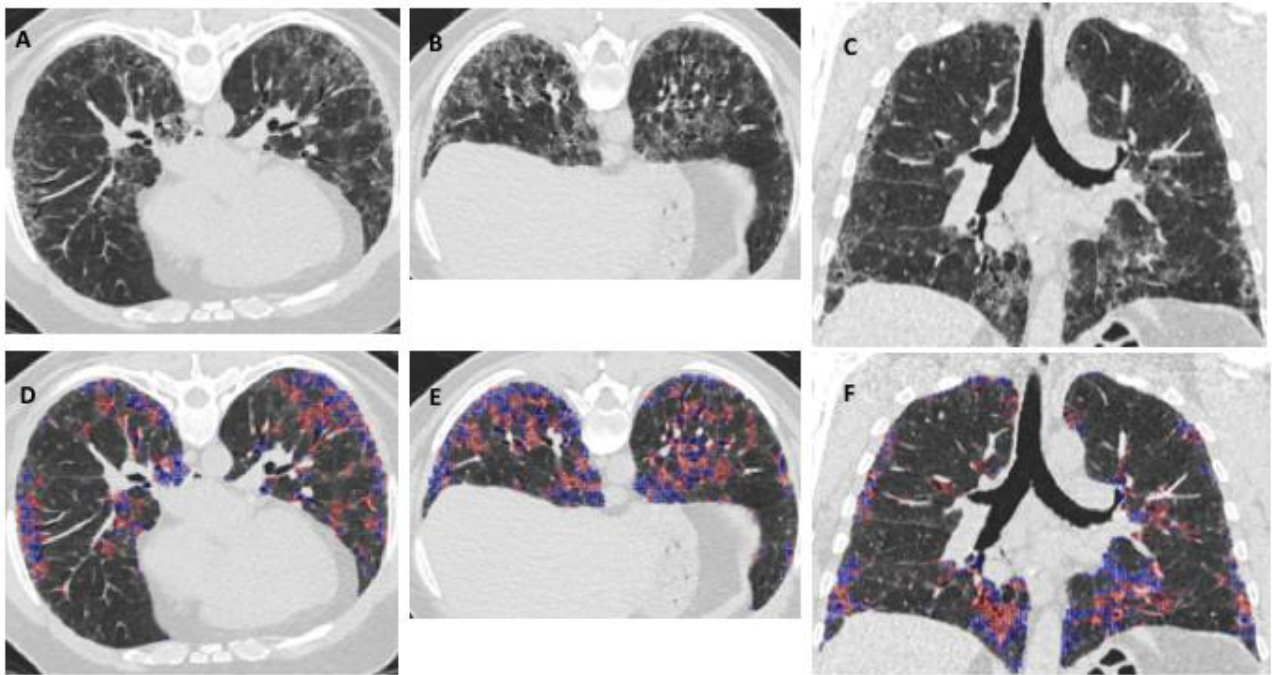


Figure 3.5 A 57-year-old male diagnosed with idiopathic pulmonary fibrosis, who had a forced vital capacity of 69% percent predicted at baseline. Axial CT images of the middle (A) and lower (B) lungs are demonstrated with coronal reconstructions (C) show peripheral reticulation and traction bronchiectasis in keeping with a probable usual interstitial pneumonia pattern. Corresponding image overlays following classification of quantitative lung fibrosis (QLF) score are also shown (D-F). The sum of blue and red dots indicate the voxels that have been classified as fibrotic reticulation. (Courtesy of Grace Hyun Kim, Center for Computer Vision and Imaging Biomarkers, Department of Radiological Sciences, David Geffen School of Medicine at UCLA, Los Angeles, USA).



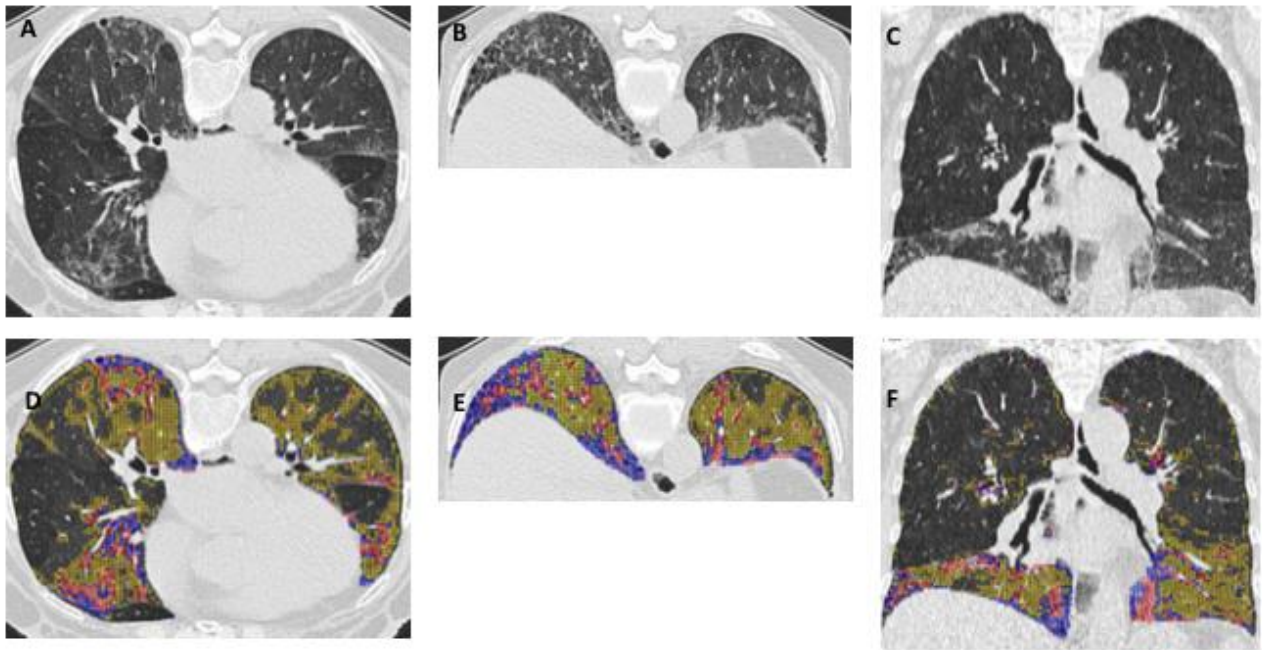


Figure 3.6 A 62-year-old man diagnosed with rheumatoid arthritis-related interstitial lung disease, who had a forced vital capacity of 77% percent predicted at baseline. Axial CT images of the middle (A) and lower (B) lungs are demonstrated with coronal reconstructions (C) which highlight peripheral reticulation and traction bronchiectasis. Corresponding image overlays following classification of quantitative interstitial lung disease (QILD) score are also shown (D-F). The sum of blue and red dots indicate the voxels that have been classified as fibrotic reticulation. The sum of yellow dots represents quantitative ground glass (QGG) lung. The sum of pink dots represents quantitative honeycomb (QHC) lung. The total sum of QLF, QGG and QHC constitutes the QILD score. (Courtesy of Grace Hyun Kim, Center for Computer Vision and Imaging Biomarkers, Department of Radiological Sciences, David Geffen School of Medicine at UCLA, Los Angeles, USA).

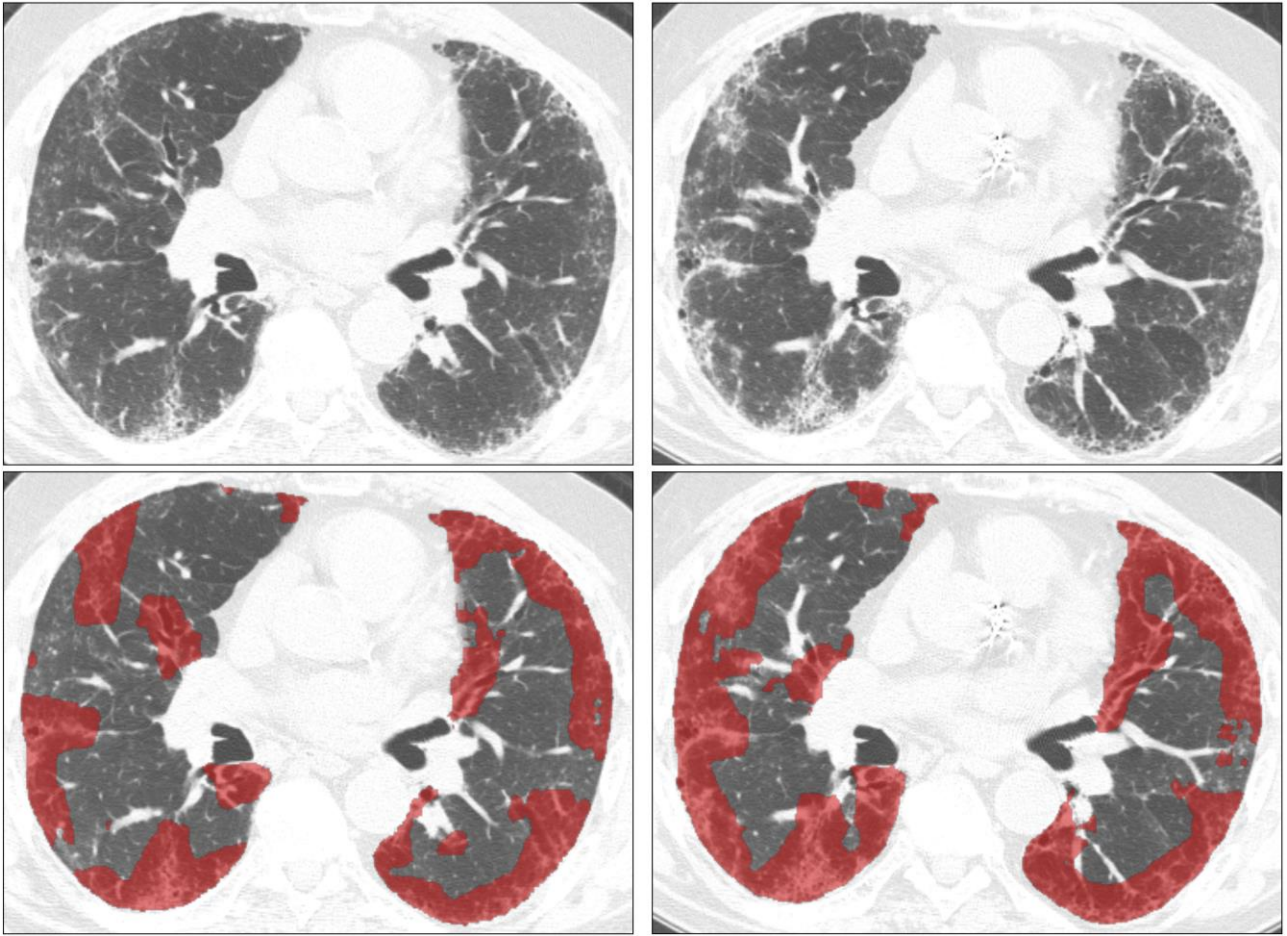


Figure 3.7 Sequential quantitative CT measurement of fibrosis in a 71-year-old male patient diagnosed with idiopathic pulmonary fibrosis. Axial CT images are shown at baseline (left) and 14 months later (right), using Data-Driven Texture Analysis (DTA). The extent of fibrosis, shown in red (bottom row), increased from 39% (bottom left) to 52% (bottom right), and the patient deteriorated physiologically, with a decrease in forced vital capacity from 66% predicted to 45% and a decrease in diffusion capacity of carbon monoxide from 43% to 26%. (Courtesy of Stephen Humphries, Quantitative Imaging Laboratory, Department of Radiology, National Jewish Hospital, Denver, Colorado, USA).